Enhanced Technique for Credit Card Extortion Detection Using Extreme Gradient Boosting Algorithm

P. Sujitha[1*] & R. Vanitha[2]

[1]UG Student, [2]Assistant Professor, [1,2]Department of CSE, IFET College of Engineering, Villupuram, India.
*Corresponding Author (P.Sujitha) Email: sujithaarul14@gmail.com**

DOI: https://doi.org/10.46431/MEJAST.2023.6205

Article Received: 10 March 2023 *Article Accepted: 28 April 2023* *Article Published: 12 May 2023*

ABSTRACT

The most common issue in the modern world is the identification of credit card fraud. This is a result of the expansion of both online commerce platforms and online transactions. In utmost cases, credit card fraud occurs when the card is stolen and used for any unauthorised exertion, or indeed when the fraudster utilises the card's information for their own gain. The credit card scam detection system was introduced with machine learning algorithms to catch these actions. Financial fraud is a growing problem in the financial industry with long-term consequences. It becomes difficult for two main reasons: first, the profiles of legitimate and fraudulent behaviour are always changing, and second, the data sets for credit card fraud are quite biased. The main objectives of this study are to identify the various types of fraudulent credit cards and to investigate alternate fraud detection techniques. On severely skewed credit card fraud data, it evaluates the performance of Decision tree, Random Forest, Logistic Regression and Extreme Gradient Boosting (XG Boost).

Keywords: Fraud detection; Machine learning; XG Boost; Credit card frauds.

1. Introduction

Financial fraud is a problem that is getting worse and has far-reaching effects on the government, businesses, and financial sector. Credit card transactions have increased in the modern world due to a strong reliance on internet technology, yet credit card fraud has also increased both online and offline. Recent computational approaches have received attention as credit card transactions become a common form of payment for goods and services. The prevention of frauds in industries and enterprises including credit card, retail, e-commerce, insurance, and others is made possible by a variety of fraud detection tools and software. One famous and well-liked solution for addressing the issue of credit fraud detection is machine learning.

Absolute certainty regarding the genuine intention and legality of a request or transaction is unattainable. Fraud may come from a credit card that has been lost, stolen, or fraudulently fabricated. Due to the rise in online purchasing, card-not-present fraud—or the use of your credit card number in e-commerce transactions—has also increased in frequency. The growth of e-banking and various online payment environments has led to an increase in fraud, such as CCF, causing billions of dollars in losses annually. CCF detection has emerged as one of the key objectives in this era of digital payments. As a company owner, it is indisputable believe a cashless society is the way of the future. As a result, conventional payment methods won't be employed in the future and won't be useful for growing a firm. In actuality, utilizing mathematical algorithms is the most efficient way to look for potential signs of fraud in the data that is accessible. Credit card fraud detection is actually the process of classifying fraudulent transactions into two categories: legitimate transactions and fraudulent transactions. To detect credit card fraud, a number of techniques have been developed and put into use including Decision Tree, Random Forest, Logistic Regression, and Extreme Boosting Algorithm (XG Boost) which is used for comparison analysis. Datasets for credit card transactions are infrequent, wildly unbalanced, and distorted. The most crucial component of machine learning to assess the effectiveness of strategies on skewed credit card fraud data is selecting the best feature (variables) for the models and choosing the right metric. Credit card detection faces a variety of difficulties, including the fact that the profile of fraudulent behavior is dynamic, meaning that fraudulent transactions frequently resemble valid ones. The effectiveness of credit card fraud detection is significantly impacted by the choice of variables, sampling strategy, and detection method. In the end, assessments of the outcomes of classifier evaluation testing are compiled. According to the results of the trials, XG Boost has a 99.94% accuracy rate. However, when all of the classifiers' learning curves are compared, it becomes clear that XG Boost overfits while Random Forest, Logistic Regression, and Decision Tree underfit. The accuracy of XG Boost is higher than that of every algorithm. Hence we conclude that XG Boost (Extreme Gradient Boosting). Hence we conclude that XG Boost (Extreme Gradient Boosting) is the best model for our system.

2. Related Works

There have been numerous research studies in the area of credit card fraud detection. This section contains many research papers focused on detecting credit card fraud. Additionally, we emphasise significantly the research that revealed fraud detection in the issue of class imbalance. Credit cards are detected using a variety of methods. Y.Abakarim [1] enforced as a result, the primary methodologies can be categorised into areas such machine learning (ML), credit card fraud detection, ensemble and feature ranking, and user authentication approaches [1],[3]. Each of ML's various branches can handle a variety of learning tasks. However, there are various framework types for ML learning.

V.Arora [3], enforced a remedy for credit card fraud is offered by the ML technique, such as random forest (RF). The random forest is the decision tree's ensemble. The RF method is used by most studies. We can utilise network analysis and (RF) to merge the model. This approach is known as APATE [1]. Different machine learning (ML) methods, including supervised learning and unsupervised methods, are available to researchers. For CCF identification, ML techniques like LR, ANN, DT, SVM, and NB are frequently employed. To build reliable detection classifiers, the researcher can combine these strategies with ensemble techniques [3]. An artificial neural network is a collection of connected neurons and nodes. An input layer, an output layer, and one or more hidden layers are only a few of the layers that make up a feed-forward perceptron multilayer. The output layer offers the algorithm's response at the end. To reduce inaccuracy, the training set was previously used with weights in the first set. H.Abdi [2] enforced these weights were all modified using intricate methods like supervised and multilayer perceptron like backpropagation [6]. V.N.Dornadula [10], I.Benchaji [11] enforced regression problem and a support vector machine (SVM) are both used in the linear classification model. We may determine the points from both classes that are closest to the line using the SVM technique. The integration of supervised and unsupervised techniques for the classification of credit card fraud detection is the focus of this research. K.Kirasic [7] on research of Random forest vs logistic regression, compared the analysis of RF and LR used Binary classification for heterogeneous settings. Khatri et al. [9] enforced several ML algorithms for credit card fraud discovery. In this exploration, the authors enforced the ensuing styles Decision Tree (DT), k-Nearest Neighbor (kNN), Logistic Retrogression (LR), Random Forest (RF), and Naive Bayes (NB). To estimate the ML-grounded credit card fraud discovery models, the experimenters used a dataset that was generated from European cardholders in 2013 also, the authors considered the perceptivity and the perfection as the main performance criteria. The results showed that the kNN algorithm achieved the most optimal results with a perfection of 91.11 and a perceptivity of 81.19. Rajora et al. [10] conducted a relative exploration of ML styles for credit card fraud discovery using the European cardholders dataset. Some of the styles that were delved include the RF and the kNN styles. The authors considered the accuracy and the area under the curve (AUC) as the main performance criteria. The results demonstrated that RF algorithm achieved an accuracy of and a AUC of 0.94. In discrepancy, the kNN obtained an accuracy of 93.2 and a AUC of 0.93. Although these results are promising, this exploration didn't probe the class imbalance issue that exists in the dataset that was used. Trivedi et al. [11] proposed an effective credit card fraud discovery machine using ML styles. In this exploration, the authors considered numerous supervised ML ways including grade Boosting (GB) and Random Forest (RF). The authors estimated these styles using the European cardholders dataset. The performance criteria used to assess the effectiveness of the proposed approaches include the accuracy and the perfection. The outgrowth of the trials showed that the GB attained an accuracy of 94.01 and a perfection of 93.99. On the other hand, the RF achieved an accuracy of 94.00 and a perfection of 95.98.

Tanouz et al. [12] presented a credit card fraud discovery frame using ML algorithms. In this exploration, the authors used the European cardholders dataset to assess the performance of the proposed styles. also, the authors enforced an under- slice fashion to break the issue of class imbalance that live in the dataset that was used. The ML styles considered in this work include the RF and LR. The experimenters used the accuracy as the main performance metric. The results demonstrated that the RF approach achieved a fraud discovery accuracy of 91.24. In discrepancy, the LR method attained an accuracy of 95.16. Likewise, the authors reckoned the confusion matrix to assert whether these proposed styles performed optimally for the positive and negative classes. The results showed that the class imbalance issue that live in the European credit card holder dataset requires farther disquisition.

Riffi et al. [13] enforced a credit card fraud discovery machine using the Extreme Learning Machine(ELM) and Multilayer Perceptron (MLP) algorithms. Both the ELM and MLP are artificial neural networks(ANNs); still, they differ in terms of internal armature. The authors used the fraud discovery accuracy as the main performance metric. The results demonstrated that the MLP system achieved an accuracy of 97.84. In discrepancy, the ELM attained credit card fraud discovery accuracy of 95.46. This work concluded that the MLP outperformed the ELM; still, the ELM is less complex in comparison to the MLP.

Randhawa et al. [14] proposed a credit card fraud discovery machine using Adaptive Boosting and Majority Voting styles. In this exploration, the authors used the European cardholders dataset. also, the authors considered the AdaBoost method in convergences with ML styles similar as the Support Vector Machine (SVM). In the trials, the accuracy and the Matthews Correlation Measure (MCC) were considered as the main performance criteria. The outcomes demonstrated that the AdaBoost- SVM achieved an accuracy of 99.85 and a MCC of 0.044. Fawaz Khaled Alarfaj [15] proposed a Credit Card Fraud Detection Using State- of- the- Art Machine Learning and Deep Learning Algorithms. In this exploration ML styles of Extreme Learning Method, Decision Tree, Random Forest, Support Vector Machine, Logistic Regression and XGBoost. The results demonstrated that the accuracy XG Boost is 99.92%.

3. Proposed System

In this section, the proposed techniques used for detecting the frauds in credit card system is to the find the fraud transaction of a credit card for the given dataset. The algorithm used for the classification of a dataset of a fraud and non-fraud transaction such as Logistic Regression, Decision Tree, Random Forest and Extreme Gradient Boosting (XG Boost). Principal Component Analysis Dimensionality reduction is used to protect user identities. The Performance of the techniques is evaluated based on precision, recall, f1 score, support, Accuracy, Sensitivity, Specificity and ROC Curve is created for showing the performance of a classification. The result that has been concluded is that Logistic regression has an accuracy of 99.89% while Decision tree shows accuracy of 99.87%, and Random forest shows accuracy of 99.88% but the best results are obtained by Extreme Boosting Algorithm (XG Boost) with a precise accuracy of 99.94%.

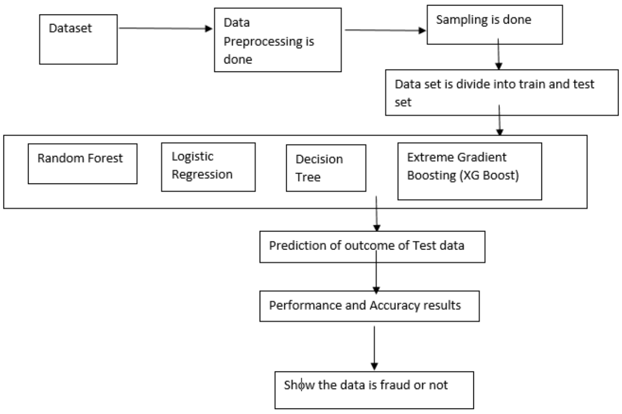

Dataset
Data Preprocessing is done
Sampling is done
Data set is divide into train and test set
Random Forest
Logistic Regression
Decision Tree
Extreme Gradient Boosting (XG Boost)
Prediction of outcome of Test data
Performance and Accuracy results
Show the data is fraud or not

Figure 1. Block Diagram

(i) Dataset Description

The well-known credit card fraud detection dataset is used in this study. The dataset includes credit card transactions made by clients of credit cards. Just 492 out of 284 807 transactions were fraudulent, creating an unbalanced dataset. Due to the change performed on the dataset, all attributes other than "Time" and "Amount" are numerical. For confidentiality purposes, these attributes are classified as V1, V2,… ,V28. The "Amount" field represents the transaction's cost, while the "Time" attribute represents the number of seconds that passed between a transaction and the dataset's initial transaction. The dependent variable, or attribute "Class," has a value of 1 for fraudulent transactions and 0 for lawful transactions.

S.No.	Feature	Description
1.	Time	Time in seconds to require the lapses between the current transaction and the first transaction
2.	V1, V2, V3……V28 attributes	These 28 columns show result of a PCA dimensionality reduction to protect user identities and sensitive features

3.	Amount	Amount of transaction
4.	Class label	Binary class labels 1 and 0 for nonfraudulent and fraudulent

Applied Machine Learning Techniques

(a) Logistic Regression

An outcome that has two possible values, such as zero or one, no or yes, false or true, is predicted by the supervised classification method known as logistic regression. Logistic regression returns the probability of a binary dependent variable that is predicted from the independent variable of the dataset. While logistic regression and linear regression have many similarities, logistic regression yields a curve as opposed to linear regression's straight line. Based on the use of one or more predictors or independent variables, logistic regression generates logistic curves that plot values between 0 and 1.

$$p = \frac{e^{\alpha+\beta_n X}}{1 + e^{\alpha+\beta_n X}}$$

(b) Decision Tree

An algorithm known as a decision tree employs conditional control statements to forecast the ultimate decision using a tree-like graph or model of decisions and their potential outcomes. A learnt function is used to represent a decision tree, which is a technique for approaching discrete-valued target functions. These kinds of algorithms are well known for inductive learning and have been effectively used for a variety of tasks. A new transaction is given a label to indicate whether it is legitimate or fraudulent; the transaction value is then checked against the decision tree, and finally, a path is shown from the root node to the transaction's output or class label.

$$Entropy(S) = \sum_{i=1} -p_i \log_2 p_i$$

$$Gain(S,A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v)$$

(c) Random Forest

An approach for classification and regression is called Random Forest. In a nutshell, it is a group of decision tree classifiers. As it corrects the inclination of overfitting to their training set, random forest has an advantage over decision trees. To train each individual tree, a subset of the training set is randomly taken, and after a decision tree has been formed, each node is divided on a feature chosen at random from the whole feature set. Because each tree is trained independently of the others in a random forest, training is incredibly quick even for big data sets with numerous characteristics and data occurrences.

$$\frac{1}{X} \sum_{x=1}^{x} f_x(\dot{R})$$

(d) XG BOOST-Extreme Gradient Boosting

The proposed XG Boost model stands for Extreme Gradient Boosting. Boosting involves several steps. In order to enhance the prediction in following rounds, several trees are formed, with the information from the first tree being supplied as input to the second tree. In essence, it is an additive tree model where new trees are added to complete the ones that have previously been constructed. XG Boost only functions with numeric data and accommodates missing values. XG Boost is a distributed gradient boosting toolkit that has been tuned for quick and scalable machine learning model training. A number of weak models' predictions are combined using this ensemble learning technique to get a stronger prediction. Extreme Gradient Boosting is one of the most well-known and commonly used machine learning algorithms because it can handle enormous datasets and perform at the cutting edge in many machine learning tasks including classification and regression .Its effective handling of missing values, which enables it to handle real-world data with missing values without requiring a lot of pre-processing, is one of the key characteristics of XG Boost. Moreover, XG Boost includes integrated parallel processing capability, allowing you to train models on huge datasets quickly.

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \epsilon \mathcal{F}$$

$$obj(\theta) = \sum_i^n l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$$

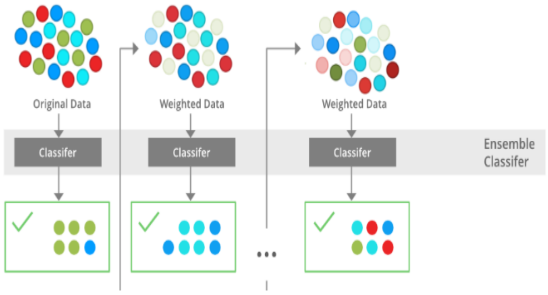


Figure 2. Structure of XG Boost

(e) Performance-Evaluation Measures

(i) Accuracy

Accuracy is used to measure the performance in the evidence domain recovery and processing of the data. The fraction of the results that are successfully classified can be represented by equation as follows:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}$$

(ii) Precision

Precision is a performance assessment that measures the ratio of correctly identified positives and the total number of identified positives. This can be seen as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

(iii) F-Measure/F1-Score

The f-measure considers both the precision and the recall. The f-measure may be assumed to be the average weight of all values, which can be seen as follows:

$$F = \frac{2 \times \text{precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

(iv) Recall

The recall is also referred to as the sensitivity, which is the ratio of connected instances retrieved over the total number of retrieved instances and can be seen as follows:

$$\text{Recall} = \frac{TP}{TP + FN}$$

4. Experimental Results

(a) Data Visualisation

In the experiment, 90% of the data was utilized as the training set while the remaining 10% served as the test set. The dataset includes 492 frauds out of 807 deals. It covers only fine input variables, which are the outgrowth of a PCA metamorphosis. Due to the issue of concealment, we cannot offer the structures of the original dataset and the data more background information. The point Time' covers the seconds ceased between the first transaction in the dataset and each transaction. Figure 3 shows the class distribution of the Credit Card Fraud dataset into a fraudulent and non-fraud transactions.

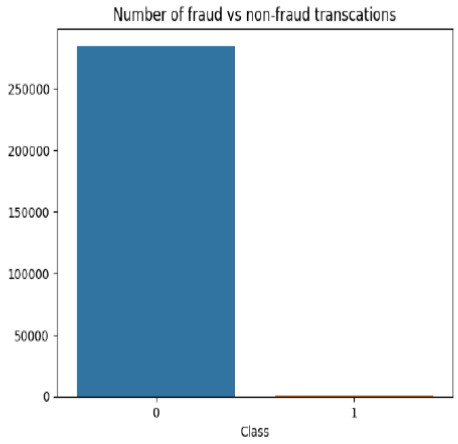


Figure 3. Fraud Vs Non Fraud

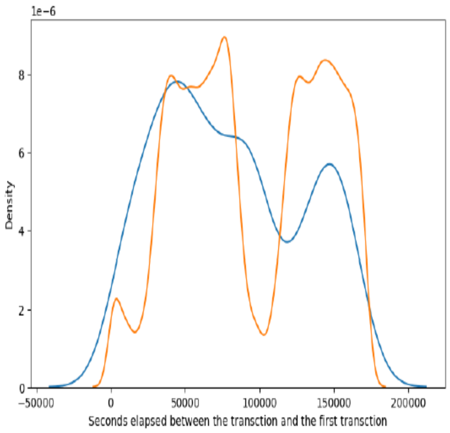


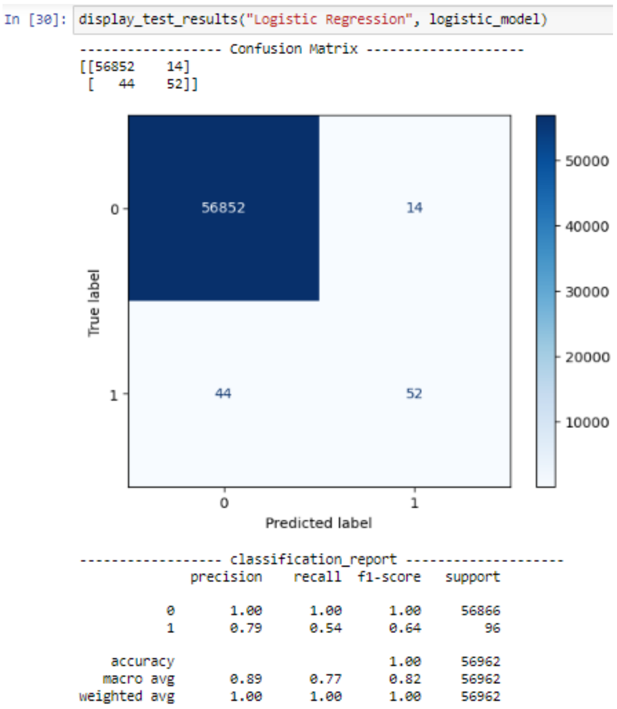
Figure 4. Distribution of class with time

(b) Accuracy of Machine learning algorithm

S.No.	Algorithm Name	Accuracy Score (%)	F1 Score (%)
1.	Decision Tree	99.87	61.53
2.	Random Forest	99.88	64.4
3.	Logistic Regression	99.89	64.19
4.	XG Boost	99.94	82.28

The data is analyzed and the behavior and pattern of the dataset and draw the features for further testing and training. Here used the labelled dataset. Finally, data are trained using the Extreme Gradient Boosting Algorithm (XGBOOST).

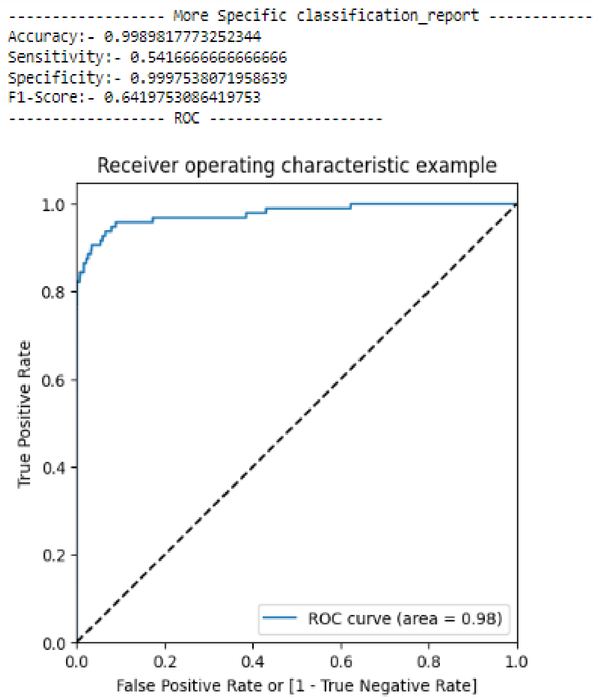
(1) Logistic Regression Result



```
In [30]: display_test_results("Logistic Regression", logistic_model)

         ----------------- Confusion Matrix -------------------
         [[56852    14]
          [   44    52]]
```

```
         ----------------- classification_report -------------------
                       precision    recall  f1-score   support

                    0       1.00      1.00      1.00     56866
                    1       0.79      0.54      0.64        96

             accuracy                           1.00     56962
            macro avg       0.89      0.77      0.82     56962
         weighted avg       1.00      1.00      1.00     56962
```

Figure 5.1. Logistic regression

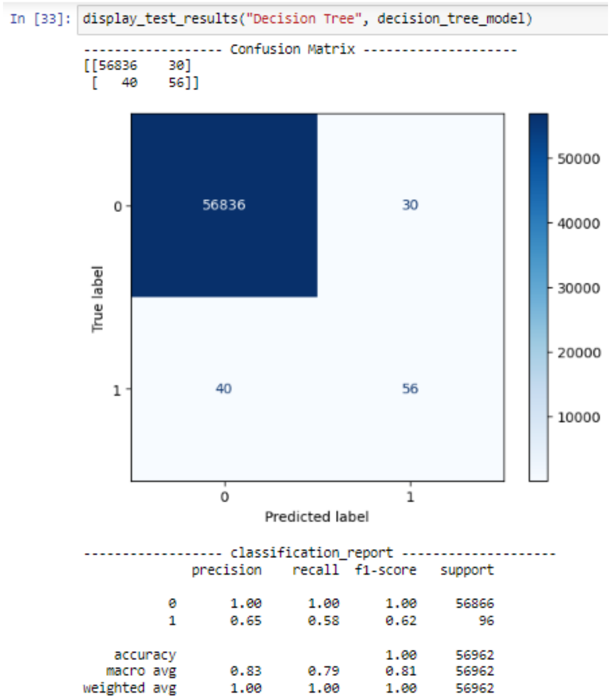


```
----------------- More Specific classification_report -----------
Accuracy:- 0.9989817773252344
Sensitivity:- 0.5416666666666666
Specificity:- 0.9997538071958639
F1-Score:- 0.6419753086419753
----------------- ROC -------------------
```

Receiver operating characteristic example
1.0
0.8
0.6
0.4
0.2
0.0
True Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
False Positive Rate or [1 - True Negative Rate]
ROC curve (area = 0.98)

Figure 5.2. Logistic regression ROC Curve

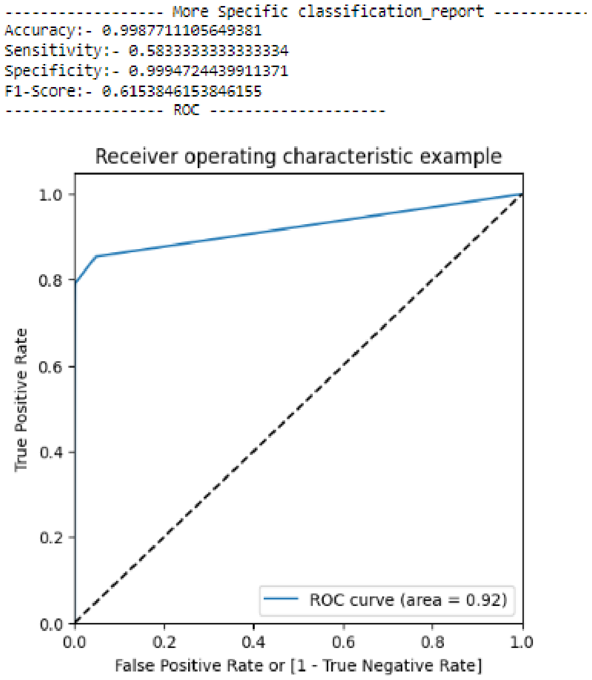
(2) Decision Tree Result



```
In [33]: display_test_results("Decision Tree", decision_tree_model)

         ----------------- Confusion Matrix -------------------
         [[56836    30]
          [   40    56]]
```

```
         ----------------- classification_report -------------------
                       precision    recall  f1-score   support

                    0       1.00      1.00      1.00     56866
                    1       0.65      0.58      0.62        96

             accuracy                           1.00     56962
            macro avg       0.83      0.79      0.81     56962
         weighted avg       1.00      1.00      1.00     56962
```

Figure 6.1. Decision tree

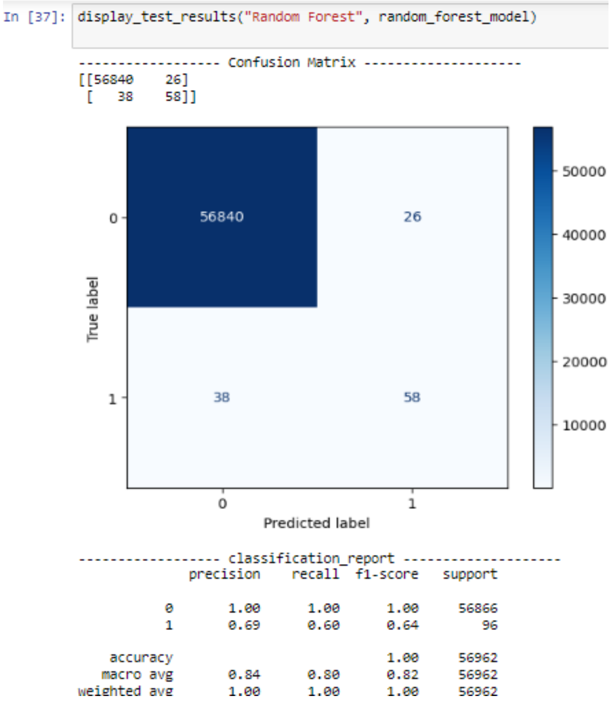


```
----------------- More Specific classification_report ----------
Accuracy:- 0.9987711105649381
Sensitivity:- 0.5833333333333334
Specificity:- 0.9994724439911371
F1-Score:- 0.6153846153846155
----------------- ROC -------------------
```

Receiver operating characteristic example
1.0
0.8
0.6
0.4
0.2
0.0
True Positive Rate
0.0
0.2
0.4
0.6
0.8
1.0
False Positive Rate or [1 - True Negative Rate]
ROC curve (area = 0.92)

Figure 6.2. Decision tree ROC Curve

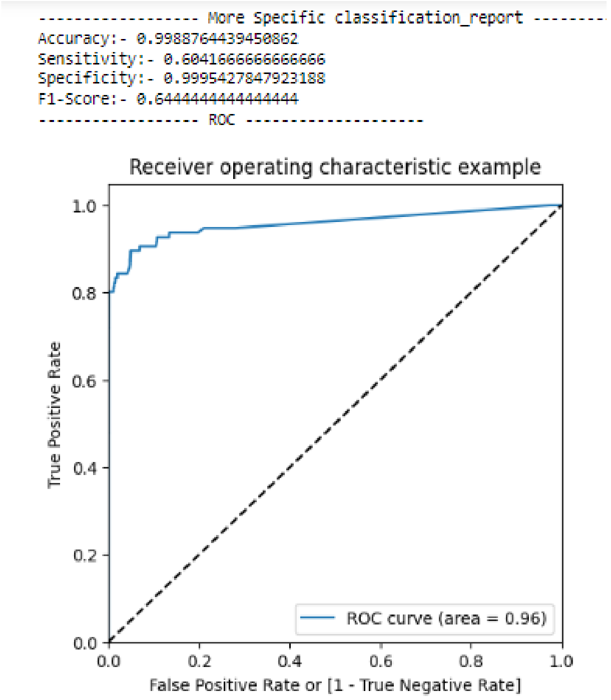
(3) Random Forest Result



```
In [37]: display_test_results("Random Forest", random_forest_model)

         ------------------ Confusion Matrix --------------------
         [[56840    26]
          [   38    58]]
```

```
         ------------------ classification_report --------------------
                       precision    recall  f1-score   support

                    0       1.00      1.00      1.00     56866
                    1       0.69      0.60      0.64        96

             accuracy                           1.00     56962
            macro avg       0.84      0.80      0.82     56962
         weighted avg       1.00      1.00      1.00     56962
```

Figure 7.1. Random Forest



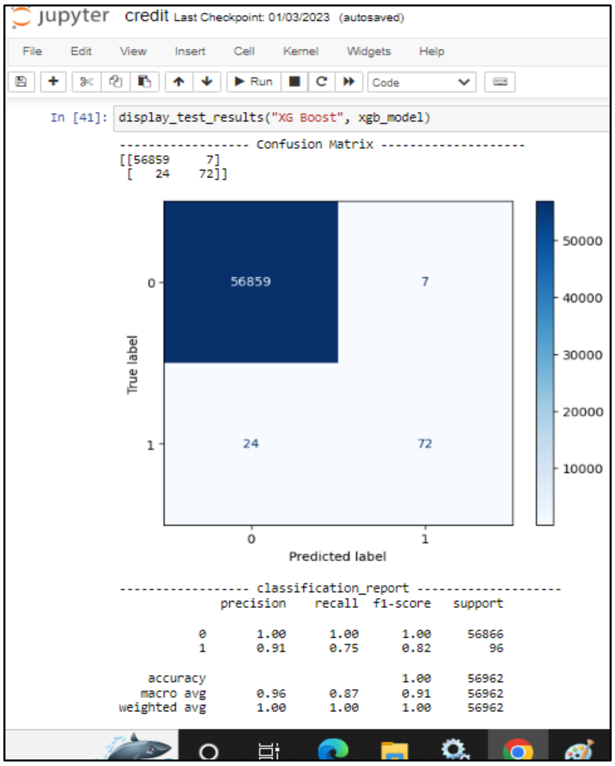
```
------------------ More Specific classification_report ---------
Accuracy:- 0.9988764439450862
Sensitivity:- 0.6041666666666666
Specificity:- 0.9995427847923188
F1-Score:- 0.6444444444444444
------------------ ROC --------------------
```

Receiver operating characteristic example
True Positive Rate
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
False Positive Rate or [1 - True Negative Rate]
ROC curve (area = 0.96)

Figure 7.2. Random Forest ROC Curve

(4) XG BOOST Result

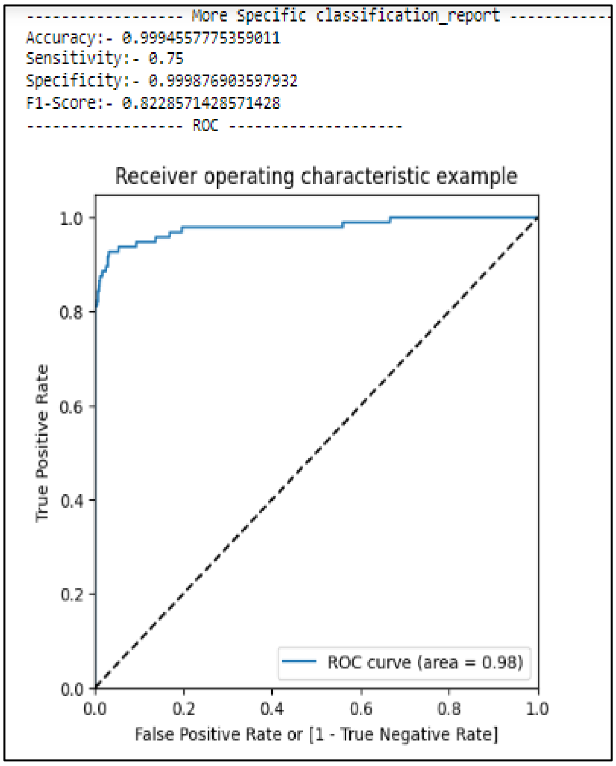
The result occurred in XG Boost accuracy is 99.94.and the f1-score is 82.28. XG Boost efficient way for finding the credit card fraud detection is best.



```
In [41]: display_test_results("XG Boost", xgb_model)

         ------------------ Confusion Matrix --------------------
         [[56859     7]
          [   24    72]]
```

```
         ------------------ classification_report --------------------
                       precision    recall  f1-score   support

                    0       1.00      1.00      1.00     56866
                    1       0.91      0.75      0.82        96

             accuracy                           1.00     56962
            macro avg       0.96      0.87      0.91     56962
         weighted avg       1.00      1.00      1.00     56962
```

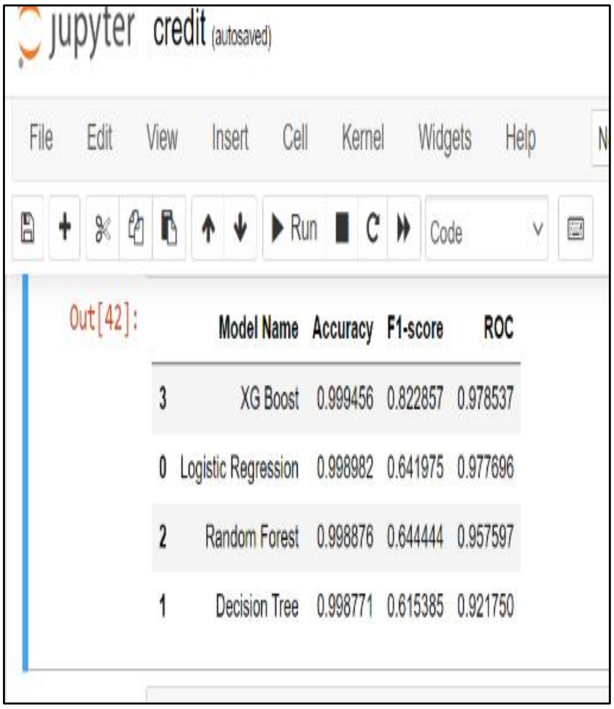
Figure 8.1. XG Boost



```
------------------ More Specific classification_report -----------
Accuracy:- 0.9994557775359011
Sensitivity:- 0.75
Specificity:- 0.999876903597932
F1-Score:- 0.8228571428571428
------------------ ROC --------------------
```

Receiver operating characteristic example
True Positive Rate
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
False Positive Rate or [1 - True Negative Rate]
ROC curve (area = 0.98)

Figure 8.2. XG Boost ROC Curve



Out[42]:

	Model Name	Accuracy	F1-score	ROC
3	XG Boost	0.999456	0.822857	0.978537
0	Logistic Regression	0.998982	0.641975	0.977696
2	Random Forest	0.998876	0.644444	0.957597
1	Decision Tree	0.998771	0.615385	0.921750

Figure 9. Comparison of algorithm

5. Conclusions and Future Recommendations

In this paper, Machine learning technique like Logistic regression, Decision Tree, Random forest, and XG Boost classifiers were used to detect the fraud in credit card system. Sensitivity, Specificity, accuracy and f1-score are used to evaluate the performance for the proposed system. From the experiments, the result that has been concluded is that Logistic regression has an accuracy of 99.89% while Decision tree shows accuracy of 99.87% and Random forest shows accuracy of 98.88% but the best results are obtained by XG Boost with a precise accuracy of 94.94%. However when the learning curves of all the classifiers are evaluated, we see that XG Boost overfits along with Random forest and decision tree. Hence we conclude that Extreme gradient boosting algorithm (XG Boost) is the best model for detecting credit card fraud detection. In future, the fraud details will be send for the respective card owners through email or message.

Declarations

Source of Funding

This study did not receive any grant from funding agencies in the public or not-for-profit sectors.

Competing Interests Statement

The authors have declared no competing interests.

Consent for Publication

The authors declare that they consented to the publication of this study.

Authors’ Contribution

Both the authors took part in literature review, research, experimentations and manuscript writing equally.

References

[1] Y.Abakarim, M. Lahby, and A. Attioui (2018). An efficient real time model for credit card fraud detection based on deep learning. In Proc. 12th Int. Conf. Intell., Pages 350-351.

[2] H. Abdi and L.J. Williams (2010). Principal component analysis. Wiley Interdiscipl. Rev. Comput. Statist., 2(4): 433-459.

[3] V. Arora, R.S. Leekha, K. Lee, and A. Kataria (2020). Facilitating user authorization from imbalanced data logs of credit cards using artificial intelligence. Mobile Inf. Syst., Pages 1-13.

[4] A. Thennakoon, C. Bhagyani, S. Premadasa, S. Mihiranga, and N. Kuruwitaarachchi (2019). Real-time credit card fraud detection using machine learning, Pages 488-493.

[5] S.P. Maniraj, A. Saini, S. Ahmed, and S. Sarkar (2021). Credit card fraud detection using machine learning and data science. Int. J. Res. Appl. Sci. Eng. Technol., 8(9): 3788-3792.

[6] J. Baszczyski, A.T. de Almeida Filho, A. Matuszyk, M. Szelg, and R. Sowiski (2021). Auto loan fraud detection using dominance-based rough set approach versus machine learning methods. Expert Syst. Appl., 163.

[7] K. Kirasich, T. Smith, and B. Sadler (2018). Random forest vs logistic regression: Binary classification for heterogeneous datasets. SMU Data Sci. Rev., 1(3): 9.

[8] N. Kousika, G. Vishali, S. Sunandhana, and M.A. Vijay (2021). Machine learning based fraud analysis and detection system. J. Phys., Conf., Number 1.

[9] S. Khatri, A. Arora, and A.P. Agrawal (2020). Supervised machine learning algorithms for credit card fraud detection: A comparison. In Proc.10th Int. Conf. Cloud Comput., Data Sci. Eng., Pages 680-683.

[10] S. Rajora, D.L. Li, C. Jha, N. Bharill, O.P. Patel, S. Joshi, D. Puthal, and M. Prasad (2018). A comparative study of machine learning techniques for credit card fraud detection based on time variance. In Proc. IEEE Symp. Comput. Intell., Pages 1958-1963.

[11] N.K. Trivedi, S. Simaiya, U.K. Lilhore, and S.K. Sharma (2020). An efficient credit card fraud detection model based on machine learning methods. Int. J. Adv. Sci. Technol., 29(5): 3414-3424.

[12] R. Sailusha, V. Gnaneswar, R. Ramesh, and G.R. Rao (2020). Credit card fraud detection using machine learning. In Proc. 4th Int. Conf. Intell. Comput. Control Syst., Pages 967-972.

[13] F.Z. El Hlouli, J. Rif, M.A. Mahraz, A. El Yahyaouy, and H. Tairi (2020). Credit card fraud detection based on and extreme learning machine architectures. In Proc. Int. Conf. Intell. Syst. Comput. Vis., Pages 1-5.

[14] K.Randhawa, C.K. Loo, M. Seera, C.P. Lim, and A.K. Nandi (2018). Credit card fraud detection using AdaBoost and majority voting, 6: 14277-14284.

[15] Fawaz Khaled Alarfaj, Iqra Malik, Hikmar Ullah Khan, et al. (2022). Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms, 17: 15-20.

[16] Altyeb Altaher Taha (2020). An Intelligent Approach to credit card fraud detection using Optimized light boosting Machine, Volume 8.